

BEST AVAILABLE COPY

PCT

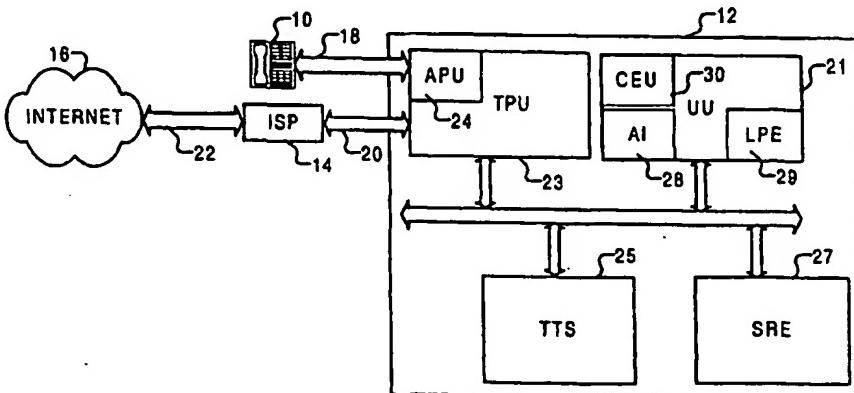
WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : H04M 3/493	A1	(11) International Publication Number: WO 00/52914 (43) International Publication Date: 8 September 2000 (08.09.00)
(21) International Application Number: PCT/US00/04587		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SL, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
(22) International Filing Date: 23 February 2000 (23.02.00)		
(30) Priority Data: 60/121,981 27 February 1999 (27.02.99) US 09/337,391 23 June 1999 (23.06.99) US		
(71)(72) Applicant and Inventor: KHAN, Emdadur, R. [US/US]; 5942 Foligno Way, San Jose, CA 95138 (US).		
(74) Agents: DALLA VALLE, Mark, A. et al.; Limbach & Limbach L.L.P., 2001 Ferry Building, San Francisco, CA 94111 (US).		Published <i>With international search report.</i>

(54) Title: SYSTEM AND METHOD FOR INTERNET AUDIO BROWSING USING A STANDARD TELEPHONE



(57) Abstract

A method and apparatus for accessing Internet using voice and audio instead of a conventional visual display. POTS (Plain Old Telephone Service) can be used to access the Internet by calling an "audio" ISP (Internet Service Provider) and interacting with an Intelligent Agent. An Audio ISP uses a standard telephone (POTS, digital or analog cellular telephone, PCS telephone, satellite telephone, etc.) instead of a modem, telephone line and traditional data ISP. The Intelligent Agent (IA) takes information from the caller, accesses the Internet, retrieves the desired information and reads it back to the caller using a voice signal. The IA can surf the net by responsively interacting with the caller using voice. The IA does not need a web browser. The IA does not require any change in the current world wide web data format to support audio. The IA works with the existing web data format. Users can also access email (both send and receive) by talking and listening through the IA using POTS.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslavia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CN	Cameroun	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LJ	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Description

5

10

15

20

25

30

35

40

45

50

55

5

5 SYSTEM AND METHOD FOR INTERNET AUDIO BROWSING USING A STANDARD
TELEPHONE

10

10

15

RELATED APPLICATIONS

The present invention claims priority to Provisional Application No.
20 15 60/121,981 filed February 27, 1999 and entitled INTERNET ACCESS USING
REGULAR PHONE. This priority document is hereby incorporated herein by
reference.

25

20

BACKGROUND OF THE INVENTION

30

Field of the Invention

The present invention relates to a method and apparatus for Internet access.
25 and more particularly to accessing and navigating the Internet through the use of
35 an audio interface via standard POTS (plain old telephone service).

Description of the related art

The number of Internet access methods has increased with the rapid growth of
40 30 the Internet. World Wide Web (WWW) "surfing" has likewise increased in
popularity. Surfing or "Internet surfing" is a term used by analogy to describe the
ease with which a user can use the waves of information flowing around the
45 Internet to find desired or useful information. The term surfing as used in this
specification is intended to encompass all of the possible activities a user can
35 participate in using the Internet. Beyond looking up a particular Internet resource or
executing a search, surfing as used herein is intended to include playing video

50

55

5 games, chatting with other users, composing web pages, reading email, applying
for an online mortgage, trading stocks, paying taxes to the Internal Revenue
Service, transferring funds via online banking, purchasing concert or airline tickets,
etc. Various kinds of web browsers have been developed to facilitate Internet
10 access and allow users to more easily surf the Internet. In a conventional web
interface, a web browser (e.g. Netscape Navigator® which is part of Netscape
Communicator® produced by Netscape Communications Corporation of Mountain
View, CA) visually displays the contents of web pages and the user interacts with
15 the browser visually via mouse clicking and keyboard commands. Thus, web
surfing using conventional web browsers requires a computer or some other an
20 Internet access appliance such as a WB-2001 WebTV® Plus Receiver produced by
Mitsubishi Digital Electronics America, Inc. of Irvine, CA.

Recently, some web browsers have added a voice based web interface in a
25 desktop environment. In such a system, a user can verbally control the visual web
browser and thus surf the Internet. The web data is read to the user by the
30 browser. However, this method of Internet access is not completely controllable by
voice commands alone. Users typically must use a mouse or a keyboard to input
35 commands and the browser only reads the parts of the web page selected using
the mouse or the keyboard. In other words, existing browsers that do allow some
40 degree of voice control still must rely on the user and visual displays to operate. In
addition, these browsers require that the web data to be read aloud must be
formatted in a specific way (e.g. the shareware Talker Plug-In written by Matt
Pallakoff and produced by MVP Solutions Inc. of Mountain View, CA can be used
45 with Netscape Commerce Server and uses files formatted in accordance with a file
50 format identified by the extension ".talk" (see i.e.
<http://www.mvpsolutions.com/PlugInSite/Talker.html> which was printed on
June 22, 1999 and is incorporated herein by reference.)

Some commercially available products (e.g. Dragon Dictate® from Dragon
Systems Inc. of Newton, MA) can read a web page as displayed on a conventional
55 30 browser in the standard web data format, however, the particular portion of the
page to be read must be selected by the user either via mouse or voice commands.
A critical limitation of these systems is that they require the user to visually examine

5 the web data and make a selection before any web data to speech conversion can
be made. This limitation also exists when using these systems to surf the web.

10 The user needs to look at the browser and visually identify the desired Uniform
Resource Locator (URL) (or use a predetermined stored list of URLs) and then

15 5 select the desired URL by voice commands. What is needed is a means to access
and surf the Internet that does not rely upon the user being able to visually perceive
web data. What is further needed is a system for "audio-only" access to the
Internet that does not require the authors of web pages to provide web data in
specialized formats for audio play-back.

SUMMARY OF THE INVENTION

In view of the background discussed above, it is an object of the present invention to provide an improved web browser interface that: does not require the use of a computer or other Internet appliance, thus making Internet access significantly simpler by using a ubiquitous device like POTS; can interact with the user completely through audio signals using voice recognition and web data to speech conversion (i.e., without any need to visually perceive web pages); and allow the use of a conventional visual browser component but with a more intelligent interface that permits audio-only control and feedback (i.e., looking at the browser is optional). Another object of the present invention is to bring Internet access to the masses of people who either cannot afford a computer or lack computer training but can use the ubiquitous POTS. Thus, the present invention allows Internet browsing without requiring the substantial cost of owning and operating a computer or Internet access appliance.

In addition, since the present invention allows a user to browse the Internet with voice only, the user is thus enabled to do so while his eyes and/or hands are otherwise occupied (e.g., while driving, walking, or operating machinery). Another object of the present invention is to facilitate audio-only web browsing using web data as currently formatted (i.e., the present invention does not require a change to

5 the existing web server data format to support audio-only browsing). Another object
of the present invention is to allow access to email using POTS.

10 Thus the present invention provides a method of browsing the Internet
comprising the steps of establishing bi-directional voice communication link with an
5 audio Internet service provider, speaking a web surfing voice command over the bi-
directional voice communication link, and then the audio Internet service provider
generating a voice response representative of a World Wide Web page
15 corresponding to the web surfing voice command. The step of generating a voice
response includes the steps of translating the spoken web surfing voice command
20 into a conventional web browser command using a speech recognition unit,
retrieving Internet data responsive to the conventional web browser command,
identifying portions of the Internet data useful to create an audio representation of
the Internet data, and translating the identified Internet data into a computer-
generated voice signal.

25 The present invention further includes a system for browsing the Internet
comprising a telephone and an audio Internet service provider coupled to the
telephone. The audio Internet service provider includes a data Internet service
provider coupled to an apparatus operable to perform a selective translation
30 function, wherein the apparatus selectively translates between voice signals and
Internet data signals. The voice signals include spoken language and the internet
data signals include World Wide Web pages. The apparatus operable to perform a
35 selective translation function includes an intelligent agent that includes a speech
recognition engine (SRE), a text to speech conversion engine (TTS), an
understanding unit (UU) for interpreting the voice signals and processing the
25 Internet data signals, and a transaction processing unit (TPU).

40 These and other features and advantages of the present invention will be
understood upon consideration of the following detailed description of the invention
and the accompanying drawings.

45

50

55

5

BRIEF DESCRIPTION OF THE DRAWINGS

10 Figure 1 depicts a high level block diagram of an example embodiment of a
5 system for accessing the Internet using a standard telephone in accordance with
the present invention.

15 Figure 2 depicts a block diagram of an example embodiment of an intelligent
agent (IA) component of the system depicted in Fig. 1 in accordance with the
present invention.

10 Figure 3 depicts a block diagram of a second example embodiment of an
intelligent agent (IA) component of the system depicted in Fig. 1 in accordance with
20 the present invention.

15 Figure 4 illustrates an example embodiment of a method of accessing the
Internet using a standard telephone in accordance with the present invention.

25

DETAILED DESCRIPTION OF THE INVENTION

30 The present invention is preferably embodied as a computer program
20 developed using an object oriented language that allows the modeling of complex
systems with modular objects to create abstractions that are representative of real
world, physical objects and their interrelationships. However, it would be
35 understood by one of ordinary skill in the art that the invention as described herein
can be implemented in many different ways using a wide range of programming
25 techniques as well as general purpose hardware systems or dedicated controllers.

40 The present invention relates to accessing the Internet using only voice and
audio instead of conventional visual inputs and displays. A POTS (plain old
telephone service) is used to access the Internet by calling an "audio" ISP (Internet
service provider). An audio ISP includes a conventional data ISP that is buffered
45 30 by an apparatus capable of performing a selective translation function using
artificial intelligence methods. In the preferred embodiment of the present
invention, this selective translation function is performed by an apparatus called an

50

55

5 Intelligent Agent (IA) which is described in detail below. The IA translates Internet
data into spoken language as well as translating spoken data and commands into
Internet web surfing commands. An audio ISP uses a standard telephone (POTS,
digital or analog cellular telephone, PCS telephone, satellite telephone, etc.)
10 instead of a modem, telephone line and a direct connection to a conventional data
ISP. An audio ISP uses TAPI (telephony application programming interface) or a
similar protocol to connect a standard telephone to a computer or other Internet
appliance. The IA takes information from the caller in the form of voice commands,
15 accesses the Internet, retrieves the desired information, and reads it back to the
caller using voice. Using voice input and output signals only, the caller can surf the
net by interacting with the IA. The IA eliminates the need for a conventional visual
20 web browser.

Turning now to FIG.1, an intelligent agent (IA) 12 allows a user, via a
standard telephone 10, to communicate with the Internet 16 through a conventional
25 ISP 14. In accordance with the present invention, the IA 12 receives voice input
signals 18 from the user via the telephone 10. One of ordinary skill in the art would
recognize that any number of audio-only-based bi-directional communication
systems could be used in place of the standard telephone 10 including digital or
30 analog cellular telephones, PCS telephones, satellite telephones, two-way radios,
etc. The IA 12 initiates an Internet session by providing a signal 20 to a
conventional ISP 14. The IA 12 can connect to the conventional ISP 14 using any
number of well known methods including the use of dial-up modems, cable
35 modems, Digital Subscriber Lines, Integrated Services Digital Networks, T1/T3
lines, Asynchronous Transfer Mode lines, local area network, high speed bus, etc.
40 The conventional ISP generates an output signal 22 to access the Internet 16 as is
known in the art. A web page from the Internet 16 is sent to the IA 12 via the
conventional ISP 14. The IA 12 interprets the contents of the web page and
determines which parts of the web page that need to be converted from text to
45 speech (TTS), text table to speech, graphics to speech (GTS), or graphics to text to
speech (GTTS using Optical Character Recognition (OCR) and then TTS). The IA
50 30 then converts the selected parts of the page to speech and sends a signal 18
containing the speech to the user via the telephone 10. The user via the telephone

5 10 can continue to request other URLs. In addition, the user can interact with web
pages such as search engines to locate a desired URL. The IA 12 repeats the
process of getting the new web page and sending back an audio-only version to the
user via the telephone 10 using, for example, a standard telephone line.

10 5 The IA 12 is configurable to provide a user-selectable level of detail in the
audio-only version of a retrieved web page. Thus, for example, a web page
containing a list of matching URLs generated by a search engine in response to a
query could be read to the user in complete detail or in summary form.

15 15 Referring now to FIG. 2, the IA 12 of Fig. 1 is described. The IA 12 provides
an intelligent interface between the user on the telephone 10 and the Internet 16.
In a basic preferred embodiment, the IA 12 includes a speech recognition engine
(SRE) 27, a text to speech conversion engine (TTS) 25, an understanding unit (UU)
21 that understands both the contents of the web page and the user's spoken
voice, and a transaction processing unit (TPU) 23. While these components of the
IA 12 are depicted as individual hardware circuits coupled together via a single bus,
one of ordinary skill in the art would understand that many different hardware
architectures could be used and likewise, the entire IA 12 (or parts of it) could be
implemented as software operable to run on a general purpose computer or even
another data processing device.

20 20 The TPU 23 communicates with the user via the telephone 10 and the
Internet 16 using signals 18 and 20. The users' telephone calls are answered by
the answer phone unit (APU) 24 which is preferably embodied as a telephone card
or modem and is part of the TPU 23. The TPU 23 communicates with the user via
the telephone 10 using, for example, the TAPI standard, a protocol developed by
25 Microsoft Corporation of Redmond, WA that is used in connecting a telephone with
a computer over a standard telephone line (see
<http://www.microsoft.com/ntserver/commserv/techdetails/prodarch/tapiwp.asp>
which was printed on June 22, 1999 and is incorporated herein by reference). In a
preferred embodiment, the TPU 23 communicates with the Internet 16 via the
40 30 conventional data ISP 14 using: a modem and a telephone line; a cable modem
and a cable line; or an Ethernet connection as is known in the art. Thus, the IA 12

5 integrates a TAPI-based audio ISP with conventional data ISP using a modem or
Ethernet connection.

10 The UU 21 is preferably implemented as a programmed computer processor
including the normally associated memory and interface ports as is well known in
the art. The UU 21 is operative to determine what part of a web page is graphics,
what part is a dynamic advertisement, what part is an interactive program, which
text is a link to a URL, etc. and makes decisions accordingly. The UU 21 is also
15 equipped with means to understand a user's commands. The UU 21 uses a
language processing engine (LPE) 29 to interpret multiple words received from the
user. The UU 21 uses an artificial intelligence (AI) unit 28 that includes one or more
20 expert systems, probabilistic reasoning systems, neural networks, fuzzy logic
systems, genetic algorithm systems, and combinations of these systems and other
systems based on other AI technologies (e.g., soft computing systems). In order to
understand the users' commands, the UU 21 uses the SRE 27 to convert users'
25 commands to text. Before sending the web page text to the user via the telephone
10, the UU 21 selectively converts text to speech using the TTS unit 25. The UU
21 allows the user to interact with Internet web pages by creating a complete audio
representation of the web pages. Thus, if a web page includes a dynamic program
30 such as a Java program to calculate a mortgage payment for example, the UU 21
20 would execute the program within the IA 12 and describe the display that would
have been generated by a conventional visual browser. The IA 12 can also use the
UU 21 to identify and interpret audio formatted data, including audio hyper-text
35 mark up language (HTML) tags.

40 The UU 21 also includes a client emulation unit (CEU) 30 that allows the UU
25 21 to execute web client type programs such as Java and Java script programs that
would normally execute on a user's client computer. The CEU 30 can spawn a
virtual machine (e.g., a Microsoft Windows NT window), execute the client program
to generate the associated displays, and pass the display data to the UU 21 to be
45 translated and relayed to the user as described above. In this way, users are able
30 to execute and interact with web pages that include executable programs.

50 FIG. 3 depicts an alternate architecture for the IA 12. The individual
functional components of the IA 12 are identical to those described in Fig. 2 and as

5 such the components are identified using the same reference numerals. The embodiment of FIG. 3 however provides a preferred arrangement for the functional components that allows a more optimized operation.

10 Turning now to FIG. 4, a flow chart depicting an example audio-only web

5 browsing transaction using the systems illustrated in FIGS. 1, 2 and 3 is described. In steps S1 and S2, a user's telephone call to the IA 12 is answered by the APU 24

15 within the TPU 23 as depicted in FIG. 2. After checking the user's identification and password in step S3, the TPU 23 asks the user for a URL to access in step S4. A connection to the conventional ISP 14 is then created in step S5 using the TPU 23.

10 After accessing the Internet and receiving the web page in step S6, the web page is interpreted by the UU 21 in step S7. In step S8, the UU 21 speaks out the appropriate text of the web page to the user via the telephone 10. Processing

steps S6 through S8 are repeated until the user discontinues selecting links to new URLs in decision step S9 and stops requesting additional URLs in decision step

15 20 25 30 35 40 45 50 55 15 At that point, the TPU 23 terminates the connections to both the telephone 10 and the internet 16.

In a preferred embodiment, the IA 12 is implemented in software and executed on a server computer. It is important to note that a user does not need a conventional visual browser because the IA 12 effectively provides an audio ISP.

20 However, the audio ISP can be implemented using a conventional visual web browser in conjunction with the IA 12. Alternatively, an audio ISP can use other means of accessing and retrieving web pages such as the Win32 Internet (Wininet) Application Programming Interface (API) as developed by Microsoft Corporation,

described at <http://pbs.mcp.com/ebooks/1575211173/ch17.htm>, printed on June

25 30 35 40 45 50 55 22, 1999 and hereby incorporated herein by reference. One of ordinary skill in the art would further understand that the IA 12 can also be used to access, manage, compose, and send email. In other words, a user can send or receive email using voice only working through the IA 12. Thus, a user can surf the web and can

exploit all of the capabilities of the Internet, simply through human voice commands and computer generated-voice responses instead of using a visual browser running on a computer or other Internet appliance.

5 While the method and apparatus of the present invention has been
described in terms of its presently preferred and alternate embodiments, those
skilled in the art will recognize that the present invention may be practiced with
modification and alteration within the spirit and scope of the appended claims. The
10 5 specifications and drawings are, accordingly, to be regarded in an illustrative rather
than a restrictive sense.

15 Further; even though only certain embodiments have been described in
detail, those having ordinary skill in the art will certainly understand that many
modifications are possible without departing from the teachings thereof. All such
10 10 modifications are intended to be encompassed within the following claims.

20

25

30

35

40

45

50

55

Claims

5

10

15

20

25

30

35

40

45

50

55

5

CLAIMS**What is Claimed is:**

- 10 5 1. A system for browsing the Internet comprising:
 a telephone; and
 an audio Internet service provider coupled to the telephone.
- 15 10 2. The system of claim 1 wherein the audio Internet service provider includes a
 data Internet service provider coupled to an apparatus operable to perform a
 selective translation function, wherein the apparatus selectively translates between
 voice signals and Internet data signals.
- 20 15 3. The system of claim 2 wherein the voice signals include spoken language and
 the internet data signals include World Wide Web pages.
- 25 20 4. The system of claim 2 wherein the apparatus operable to perform a selective
 translation function includes an intelligent agent.
- 30 25 5. The system of claim 4 wherein the intelligent agent includes at least one of a
 speech recognition engine (SRE), a text to speech conversion engine (TTS), an
 understanding unit (UU) for interpreting the voice signals and processing the
 Internet data signals, and a transaction processing unit (TPU).
- 35 35 6. The system of claim 5 wherein the UU includes a language processing engine
 (LPE) and an artificial intelligence (AI) unit.
- 40 40 7. The system of claim 5 wherein the TPU includes an answer phone unit (APU).

45

50

55

- 5 8. A system for browsing the Internet comprising:
 means for bi-directional voice communication; and
 means for providing audio Internet service coupled to the means for bi-
 directional voice communication.
- 10 5 9. The system of claim 8 wherein the means for providing audio Internet service
 includes means for providing data Internet service coupled to means for performing
 a selective translation function, wherein the means for performing a selective
 translation function is operable to selectively translate between voice signals and
 Internet data signals.
- 15 10 10. The system of claim 9 wherein the voice signals include spoken language and
 the internet data signals include World Wide Web pages.
- 20 25 11. The system of claim 9 wherein the means for performing a selective translation
 function includes at least one of means for performing speech recognition, means
 for converting text to speech, means for interpreting the voice signals and
 processing the Internet data signals, and means for processing user Internet
 surfing transactions.
- 25 30 12. The system of claim 11 wherein the means for interpreting the voice signals
 and processing the Internet data signals includes means for processing spoken
 language and means for applying artificial intelligence to determine how to
 represent and interact with a web page using only an audio signal.
- 30 35 13. The system of claim 11 wherein the means for processing user Internet surfing
 transactions includes means for responding to the initialization of a bi-directional
 voice communication.

45

50

55

- 5 14. A method of browsing the Internet comprising the steps of:
 establishing bi-directional voice communication link with an audio Internet
 service provider;
 transmitting a voice signal including a web surfing voice command over the
10 5 bi-directional voice communication link; and
 generating, by the audio Internet service provider, a voice response
 representative of an Internet data signal, the Internet data signal including a World
 Wide Web page corresponding to the web surfing voice command.
15
- 10 15. The method of claim 14 wherein the step of generating includes the step of:
 performing a selective translation function to selectively translate between
 the voice signal and the Internet data signal.
20
- 25 16. The method of claim 15 wherein the step of performing a selective translation
 function includes the steps of:
 interpreting the voice signal to identify a portion containing the web surfing
 voice command;
 performing speech recognition on the identified portion of the voice signal to
 determine the web surfing voice command;
 30 executing the web surfing voice command and receiving the Internet data
 signal in response;
 processing the Internet data signal to determine a set of user options;
 selecting text from the Internet data representative of the set of user options;
 and
 35 converting the selected text to speech.
40
- 45 17. The method of claim 16 wherein the step of processing the Internet data signal
 includes the step of applying artificial intelligence to determine how to represent
 and interact with a web page using only an audio signal, and
 30 wherein the step of interpreting the voice signal includes the step of applying
 artificial intelligence to identify the portion containing the web surfing voice
 command.

5

18. The method of claim 16 wherein the step of processing the Internet data signal includes the step of applying artificial intelligence to determine how to represent and interact with a web page using only an audio signal, and

10

5 wherein the step of performing speech recognition includes the step of applying artificial intelligence to determine the web surfing voice command.

15

19. The method of claim 14 wherein the step of establishing bi-directional voice communication link includes the step of responding to the initialization of a bi-directional voice communication.

20

20. The method of claim 14 wherein the step of generating includes the steps of:
translating the voice signal into a conventional web browser command using a speech recognition unit;

25

15 retrieving Internet data responsive to the conventional web browser command;

identifying portions of the Internet data useful to create an audio representation of the Internet data; and

30

translating the identified Internet data into a computer generated voice signal.

35

21. The method of claim 20 wherein the step of translating the voice signal includes translating a spoken email program voice control command and data, and
wherein the step of translating the identified Internet data includes the step of translating an email message into a computer generated voice signal.

40

22. The method of claim 20 wherein the step of translating the identified Internet data into a computer generated voice signal is performed by at least one of a text to speech converter, a graphics to speech converter, and a text table to speech converter.

50

- 5 23. A computer accessible medium including a computer executable program, the
program implementing a method comprising the steps of:
10 establishing a bi-directional voice communication link between a user and an
audio Internet service provider;
5 receiving a voice signal including a web surfing voice command over the bi-
directional voice communication link; and
15 generating, by the audio Internet service provider, a voice response
representative of an Internet data signal, the Internet data signal including a World
Wide Web page corresponding to the web surfing voice command.
- 10 24. The method of claim 23 wherein the step of generating includes the step of:
20 performing a selective translation function to selectively translate between
the voice signal and the Internet data signal.
- 25 15 25. The method of claim 24 wherein the step of performing a selective translation
function includes the steps of:
30 interpreting the voice signal to identify a portion containing the web surfing
voice command;
20 performing speech recognition on the identified portion of the voice signal to
determine the web surfing voice command;
35 executing the web surfing voice command and receiving the Internet data
signal in response;
processing the Internet data signal to determine a set of user options;
selecting text from the Internet data representative of the set of user options;
25 and
40 converting the selected text to speech.

45

50

55

5 26. The method of claim 25 wherein the step of processing the Internet data signal
includes the step of applying artificial intelligence to determine how to represent
and interact with a web page using only an audio signal, and
10 wherein the step of interpreting the voice signal includes the step of applying
5 artificial intelligence.

15 27. The method of claim 25 wherein the step of processing the Internet data signal
includes the step of applying artificial intelligence to determine how to represent
and interact with a web page using only an audio signal, and
10 wherein the step of performing speech recognition includes the step of
20 applying artificial intelligence.

25 28. The method of claim 23 wherein the step of establishing bi-directional voice
communication link includes the step of responding to the initialization of a bi-
directional voice communication.

30 29. The method of claim 23 wherein the step of generating includes the steps of:
 translating the web surfing voice command into a conventional web browser
command using a speech recognition unit;
35 20 retrieving Internet data responsive to the conventional web browser
command;
 identifying portions of the Internet data useful to create an audio
representation of the Internet data; and
 translating the identified Internet data into a computer generated voice
25 signal.

40 30. The method of claim 29 wherein the step of translating the web surfing voice
command includes translating a spoken email program voice control command and
data, and
45 30 wherein the step of translating the identified Internet data includes the step
of translating an email message into a computer generated voice signal.

5 31. The method of claim 29 wherein the step of translating the identified Internet
data into a computer generated voice signal is performed by at least one of a text to
speech converter, a graphics to speech converter, and a text table to speech
converter.

10 5

15

20

25

30

35

40

45

50

55

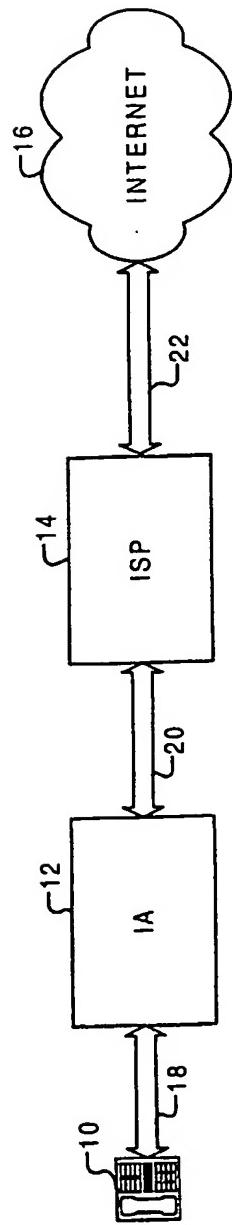


FIG. 1

2/4

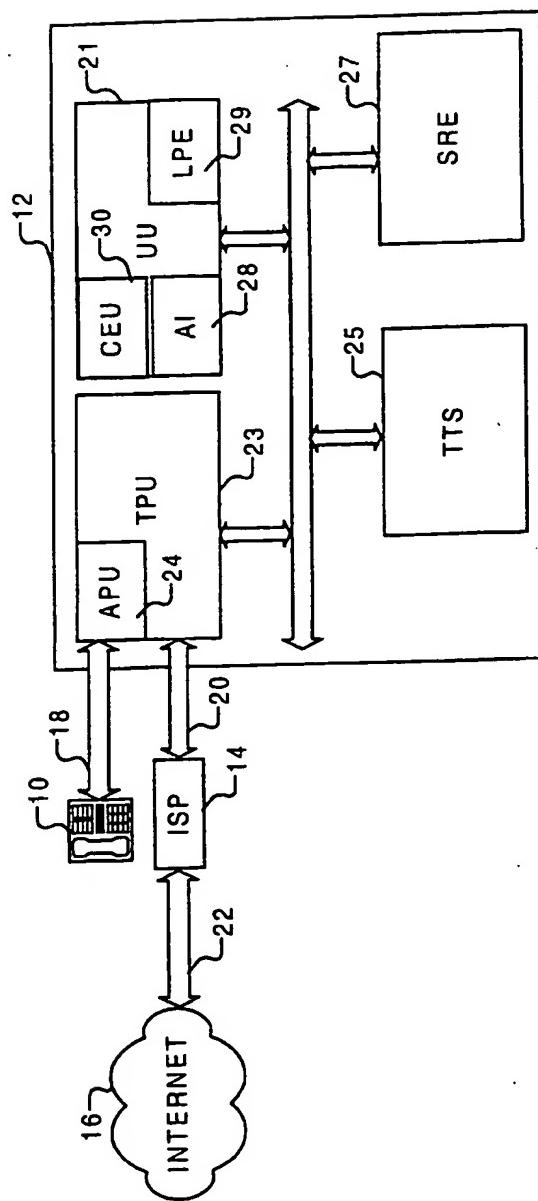


FIG. 2

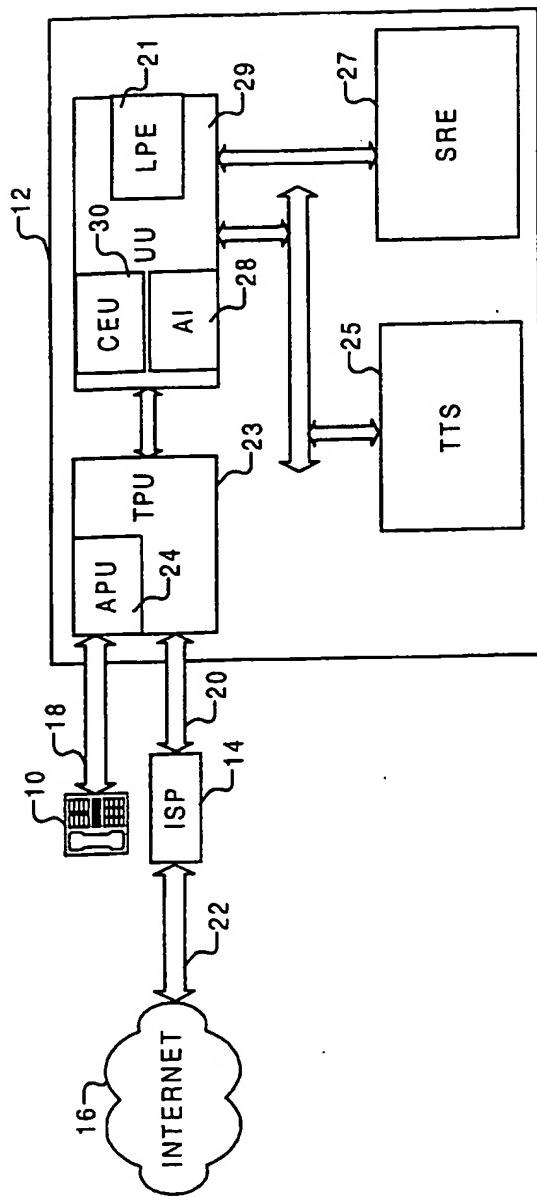
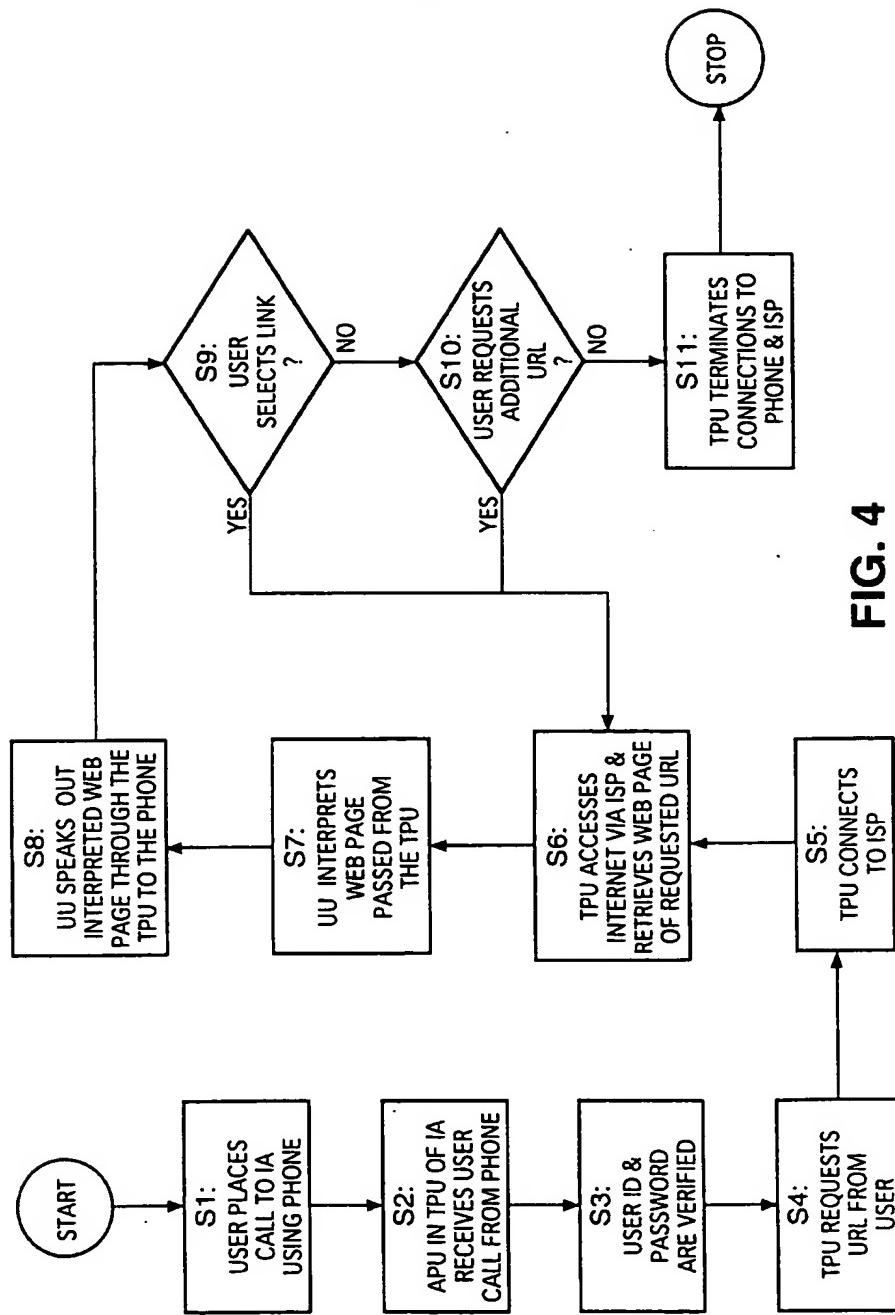


FIG. 3

4/4

**FIG. 4**

INTERNATIONAL SEARCH REPORT

Int'l. Appl. No.	PCT/US 00/04587
------------------	-----------------

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 H04M3/493

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 H04M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GB 2 307 619 A (POLLITT ALEXANDER JAMES) 28 May 1997 (1997-05-28) the whole document	1-31
X	ATKINS D L ET AL: "INTEGRATED WEB AND TELEPHONE SERVICE CREATION" BELL LABS TECHNICAL JOURNAL, US, BELL LABORATORIES, vol. 2, no. 1, 21 December 1997 (1997-12-21), pages 19-35, XP000659566 ISSN: 1089-7089 the whole document	1-20, 22-29, 31
Y	— — —/—	21, 30

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *T* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- *Z* document member of the same patent family

Date of the actual completion of the international search

9 June 2000

Date of mailing of the International search report

19/06/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5018 Patentlan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl.
Fax: (+31-70) 340-3016

Authorized officer

Megalou, M

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Ind. and Application No.
PCT/US 00/04587

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PAGE J H ET AL: "THE LAUREATE TEXT-TO-SPEECH SYSTEM - ARCHITECTURE AND APPLICATIONS" BT TECHNOLOGY JOURNAL, GB, BT LABORATORIES, vol. 14, no. 1, 1 January 1996 (1996-01-01), pages 57-67, XP000554639 ISSN: 1358-3948 page 59, paragraph 4 -page 60, paragraph 4	21, 30
X	WO 98 35491 A (JOHNSTON ROBERT DENIS ;BRITISH TELECOMM (GB)) 13 August 1998 (1998-08-13) page 1, line 24 -page 4, line 8 figures 1-3	1-20, 22-29, 31
X	GB 2 317 070 A (IBM) 11 March 1998 (1998-03-11) the whole document	1-20, 22-29, 31

INTERNATIONAL SEARCH REPORT

Information on patent family members

Title and Application No

PCT/US 00/04587

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
GB 2307619	A 28-05-1997	NONE		
WO 9835491	A 13-08-1998	AU 5674398 A EP 0958692 A		26-08-1998 24-11-1999
GB 2317070	A 11-03-1998	NONE		

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: Document in small prints.**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.